

Robust Spatio-Temporal Partial-Response Signaling over a Frequency-Selective Fading MIMO Channel with Imperfect CSI

Jelle Bailleul*, Lennert Jacobs*, Mamoun Guenach*[†], and Marc Moeneclaey*

*Department of Telecommunications and Information Processing, Ghent University
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

Email: {jelle.bailleul,lennert.jacobs,marc.moeneclaey}@ugent.be

[†]IMEC

Kapeldreef 75, 3001 Leuven, Belgium

Email: guenach@ieee.org

Abstract—Partial-response signaling is known to facilitate the equalizer design because a controlled amount of residual interference is permitted. The design of the target impulse response of the partial-response precoder often assumes perfect channel state information, which is unfortunately not available at the transmitter in most practical applications. Consequently, this contribution focuses instead on the robust and joint design of a spatio-temporal target impulse response and the equalization coefficients for a frequency-selective fading multiple-input multiple-output communication channel based on current and/or previous noisy channel estimates. More precisely, the error in the channel estimates is statistically modeled, and robustness is achieved by minimizing the mean-squared estimation error averaged over the joint distribution of the actual channel and the available channel estimates. Numerical results of the bit error rate confirm that the proposed robust partial-response signaling not only provides a significant performance gain compared to traditional full-response signaling, but also outperforms the naive approach, which ignores channel estimation errors.

I. INTRODUCTION

Compared to single-input single-output (SISO) transmission, multiple-input multiple-output (MIMO) transmission is a satisfactory solution to cope with the continuously growing demand for bit rate in modern communication systems. However, intersymbol interference (ISI) caused by the mutually interfering frequency-selective (FS) channels can considerably deteriorate the communication link quality, especially when the signal bandwidth exceeds the channel coherence bandwidth. Several techniques have therefore been proposed to counteract this unfavorable impact of the ISI. First, the optimal maximum-likelihood sequence detector (MLSD) is mostly impractical due to its excessive complexity that exponentially grows with increasing channel length [1]. Consequently, state-of-the-art transceivers envisage

more practical MIMO equalization techniques such as linear MIMO equalizers, which can be implemented at the transmitter (TX) and/or the receiver (RX). Combined with a symbol-by-symbol detector, linear equalizers serve as a suboptimal but low-complexity alternative to the MLSD. Including additional feedback from either previous decision at the RX (MIMO decision feedback equalizer [2]) or previously transmitted symbols at the TX (MIMO Tomlinson-Harashima precoding [3]) can further enhance the performance of these equalizers.

Partial-response signaling (PRS) offers a particularly interesting alternative to the traditional full-response signaling (FRS) as the former can handle large amounts of ISI [4], [5]. PRS is characterized by a target impulse response (TIR) with integer coefficients, enabling to perform the symbol-by-symbol detection in a periodically extended symbol set. This facilitates the design of the equalizer since a controlled amount of residual ISI is allowed. In [6], the joint optimization of an infinite-length linear zero-forcing equalizer and a real-valued TIR has been discussed in the context of a FS SISO channel with a whitened-matched filter RX front-end. Recently, this optimization problem has been extended to the MIMO FS channel with a finite-length, linear minimum-mean-squared error (MMSE) MIMO equalizer and a more general complex-valued PRS precoder with spatio-temporal components [7].

As the FS MIMO channel is time-variant, a new set of equalization and TIR parameters must be computed at regular intervals to ensure proper performance. Ideally, this computation is executed using perfect channel state information (CSI). However, the assumption of perfect CSI at the TX (CSIT) is often unrealistic due to imperfections such as the noise impact on the channel estimates and the limited feedback between RX and TX. In general, the naive approach of determining the

Part of the research work was conducted at the Nokia Bell Labs, Antwerp.

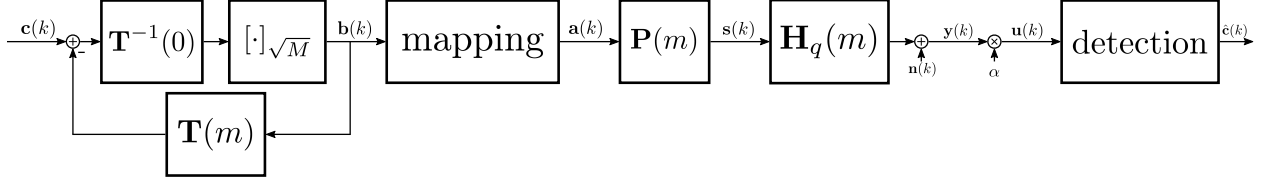


Figure 1. System model of a FS MIMO channel with a spatio-temporal MIMO PRS precoder and linear MIMO pre-equalizer at the TX.

equalization and TIR parameters without taking channel estimation errors into account yields adequate performance only when the available CSIT is sufficiently accurate. Instead, superior performance is usually attained by robust equalization designs, which incorporate the imperfections of the CSIT in the design of the equalization and TIR parameters. Two main classes of robust designs exist. As for the worst-case designs, the channel estimation error is assumed to be within a bounded uncertainty region, and a performance level is guaranteed for all errors within that region ([8], [9], [10]). Alternatively, the uncertainty of the CSIT can be modeled statistically such that a performance measure averaged over the actual channel and the available CSIT can be optimized ([11], [12]). In the sequel, the latter concept of robustness is employed.

Although the joint optimization of the TIR and equalization coefficients has been discussed in [7] for a FS MIMO channel with perfect CSIT, to our knowledge, no robust joint optimization has been presented in the literature yet for imperfect CSIT. As the performance of PRS depends greatly on the selection of the TIR, a robust design is of particular importance in practice.

In this paper, statistically robust PRS over a FS and time-variant $L \times L$ MIMO link is investigated. More precisely, the TX applies L streams of M -QAM data symbols to a spatio-temporal MIMO PRS precoder characterized by its complex-valued TIR, after which the precoded signal is equalized by a linear MIMO pre-equalizer (Section II). Both the TIR and the pre-equalizer are obtained by minimizing the mean-squared error (MSE) between the desired response and the decision variable. However, as only imperfect CSI is available, the joint optimization problem is based on the current estimate and/or previous estimates of the time-variant MIMO channel (Section III). In section IV, numerical results illustrate the performance of this robust design. Finally, conclusions are drawn in section V.

In the sequel, the notations $(\cdot)^T$, $(\cdot)^H$, $\text{Tr}(\cdot)$, and \otimes denote the transpose, the conjugate transpose, the trace, and the Kronecker product, respectively. Moreover, the all-ones column vector is denoted by $\mathbf{1}$ and all optimized variables are indicated by $(\cdot)^*$.

II. SYSTEM MODEL

This contribution considers the digital transmission over a discrete-time wide-sense stationary slowly

Rayleigh fading $L \times L$ MIMO channel. Its impulse response matrix related to the k th symbol interval is denoted by $\mathbf{H}(m, k)$, where m is the delay expressed in symbol intervals. Consecutively transmitted blocks span each N_{tot} symbol intervals, and contain both pilot symbols (for channel estimation) and data symbols. With the assumption that the channel coherence time is much larger than the block duration, $\mathbf{H}(m, k)$ is approximated by $\mathbf{H}_q(m) = \mathbf{H}(m, qN_{\text{tot}})$ during the q th block, i.e., for $qN_{\text{tot}} \leq k < (q+1)N_{\text{tot}}$. In the following, we describe the TX and RX processing related to the data symbols in a generic block, and drop the block index q for notational convenience. The channel estimation based on the pilot symbols is treated in Appendix A.

The equalized communication link is schematically depicted in Figure 1. The TX input comprises L streams of complex-valued data symbols, which are uniformly and independently drawn at a symbol rate $1/T$ from the symbol set $\mathcal{C} = \{0, 1, \dots, \sqrt{M}\} + j\{0, 1, \dots, \sqrt{M}\}$, where M is assumed to be an integer power of 4. All input symbols at the k th symbol interval are combined into the vector $\mathbf{c}(k)$, and the corresponding sequence $\{\mathbf{c}(k)\}$ related to a given block is subsequently converted by the spatio-temporal PRS precoder, described in [7], into the sequence $\{\mathbf{b}(k)\}$ according to

$$\mathbf{b}(k) = \left[\mathbf{T}^{-1}(0) \left(\mathbf{c}(k) - \sum_{m=1}^{L_T} \mathbf{T}(m) \mathbf{b}(k-m) \right) \right]_{\sqrt{M}}^{\mathbf{1}}, \quad (1)$$

where the $L \times L$ matrices $\mathbf{T}(m)$, $m \in \{0, \dots, L_T\}$ define the spatio-temporal TIR of the PRS precoder, and $[\cdot]_X$ represents the element-wise modulo- X reduction of both the real and the imaginary part to the interval $[0, X)$. Characteristic to this PRS precoder is that all elements of $\mathbf{T}(0)$ and $\mathbf{T}(m)$ are restricted to the set of Gaussian integers (which are complex numbers with integer real and imaginary parts), and that $|\det(\mathbf{T}(0))|$ is constrained to 1 (making $\mathbf{T}(0)$ unimodular). Consequently, $\mathbf{T}^{-1}(0)$ also contains only Gaussian integers and all entries of the vector $\mathbf{b}(k)$ are independently and uniformly drawn from the set \mathcal{C} [7]. The vector $\mathbf{b}(k)$ is mapped to the normalized M -QAM constellation, i.e., $\mathbf{a}(k) = 2\Delta\mathbf{b}(k) + (1+j)\Delta(-\sqrt{M}+1)\mathbf{1}$, where Δ is selected for normalizing the symbol energy ($\mathbb{E}[\mathbf{a}(k)\mathbf{a}^H(k)] = \mathbf{I}_L$). Next, the TX employs a linear, spatio-temporal MIMO pre-equalizer to combat the ISI generated by the FS channel. This pre-

equalizer has a finite-length impulse response $\{\mathbf{P}(m)\}$ with $\mathbf{P}(m) = \mathbf{0} \forall m \notin \{L_P^{(1)}, \dots, L_P^{(2)}\}$ and satisfies an energy constraint that limits the average energy per transmitted symbol, i.e.,

$$\text{Tr}(\mathbf{P}^H \mathbf{P}) = LE_{\text{TX}} \quad (2)$$

with $\mathbf{P} = [\mathbf{P}^H(-L_P^{(1)}) \dots \mathbf{P}^H(-L_P^{(2)})]^H$.

Subsequently, the pre-equalizer output signal $\mathbf{s}(k)$ is transmitted over the discrete-time MIMO channel with impulse response $\{\mathbf{H}(m)\}$. As the channel duration is essentially finite in practice, $L_H^{(1)}$ and $L_H^{(2)}$ are defined such that $\mathbf{H}(m) = \mathbf{0} \forall m \notin \{-L_H^{(1)}, \dots, L_H^{(2)}\}$, which indicates that the channel consists of $L_H = L_H^{(1)} + L_H^{(2)} + 1$ symbol-spaced taps. The channel also adds a noise vector $\mathbf{n}(k)$, whose components are spatially and temporally uncorrelated and have variance σ_n^2 .

To compensate for the constrained transmit energy per symbol, the RX scales the signal $\mathbf{y}(k)$ by a factor α before performing symbol-by-symbol detection on the resulting decision variable $\mathbf{u}(k) = \alpha \mathbf{y}(k)$, yielding

$$\mathbf{u}(k) = \alpha \sum_{m=-L_G^{(1)}}^{L_G^{(2)}} \mathbf{G}(m) \mathbf{P} \mathbf{a}(k-m) + \alpha \mathbf{n}(k), \quad (3)$$

where $L_G^{(x)} = L_H^{(x)} + L_P^{(x)}$ with $x \in \{1, 2\}$, and $\mathbf{G}(m) = [\mathbf{H}(m + L_P^{(1)}) \dots \mathbf{H}(m - L_P^{(2)})]$. As the target response vector

$$\mathbf{u}_T(k) = \sum_{m=0}^{L_T} \mathbf{T}(m) \mathbf{a}(k-m), \quad (4)$$

is related to the original data $\mathbf{c}(k)$ by

$$\mathbf{c}(k) = \left[\frac{\mathbf{u}_T(k)}{2\Delta} + \frac{(1+j)(\sqrt{M}-1)}{2} \sum \mathbf{T}(m) \mathbf{1} \right]_{\sqrt{M}}. \quad (5)$$

the decision $\hat{\mathbf{c}}(k)$ is obtained by first rounding the bracketed expression in (5), with $\mathbf{u}_T(k)$ replaced by $\mathbf{u}(k)$ from (3), to the nearest constellation point from the extended symbol set. Next, the modulo- \sqrt{M} operator reduces this decision in the extended symbol set to the original symbol set \mathcal{C}^L . Satisfactory error performance is achieved when the pre-equalizer yields a small difference between $\mathbf{u}(k)$ and $\mathbf{u}_T(k)$.

III. ROBUST MMSE EQUALIZER DESIGN

A. MMSE optimization problem

The TIR and the pre-equalizer for the q th block are optimally designed with perfect CSIT, i.e., perfect knowledge of $\mathbf{H}_q = [\mathbf{H}_q(-L_H^{(1)}), \dots, \mathbf{H}_q(L_H^{(2)})]$ at the TX. To acquire this CSIT, pilot symbols are commonly employed to estimate \mathbf{H}_q for given q at the RX, after which the corresponding estimate $\hat{\mathbf{H}}_q$ is communicated to the TX over a return channel with possibly limited bandwidth. However, this CSIT is often inaccurate

(because of channel estimation errors caused by noise) and/or delayed (because of the propagation times on the actual channel and the return channel, and the processing time needed for computing $\hat{\mathbf{H}}_q$ at the RX). This contribution discusses three particular CSIT scenarios, which differ in the available estimates at the TX: (S1) both the current estimate $\hat{\mathbf{H}}_q$ and the past K estimates $\hat{\mathbf{H}}_{q-1}, \dots, \hat{\mathbf{H}}_{q-K}$ are available; (S2) only the current estimate is available; (S3) only the past K estimates are available.

A frequently employed performance measure in the design of the TIR and the equalizer is the MSE between the decision variable $\mathbf{u}(k)$ and desired response $\mathbf{u}_T(k)$, which is defined by

$$\text{MSE} \triangleq \frac{\mathbb{E}[\|\mathbf{u}(k) - \mathbf{u}_T(k)\|^2]}{L\sigma_a^2}. \quad (6)$$

In case of perfect CSIT, \mathbf{P} and \mathbf{T} are functions of \mathbf{H}_q , so that the expectation $\mathbb{E}[\cdot]$ in (6) is over the channel impulse response \mathbf{H}_q and over all data symbols and noise samples contributing to either $\mathbf{u}(k)$ or $\mathbf{u}_T(k)$. However, \mathbf{P} and \mathbf{T} are, in practice, functions of the imperfect CSIT. In this case, the robust design of the equalization and TIR parameters takes these imperfections into account by additionally including all channel estimates used by the TX during the q th block in the expectation in (6). Because of the wide-sense stationary fading, MSE from (6) does not depend on the block index q .

Plugging (3) and (4) into (6) one obtains

$$\text{MSE} = \iint J(\mathbf{h}_q) f(\mathbf{h}_q, \hat{\mathbf{h}}_q) d\mathbf{h}_q d\hat{\mathbf{h}}_q, \quad (7)$$

where $f(\mathbf{h}_q, \hat{\mathbf{h}}_q)$ represents the joint probability density function (PDF) of the channel vector \mathbf{h}_q and the estimation vector $\hat{\mathbf{h}}_q$. The channel vector \mathbf{h}_q results from stacking the columns of \mathbf{H}_q into a single column vector with $L^2 L_H \times 1$ components. The vector $\hat{\mathbf{h}}_q$ depends on the considered CSIT scenario:

$$\hat{\mathbf{h}}_q = \begin{cases} [\hat{\mathbf{h}}_q^H \dots \hat{\mathbf{h}}_{q-K}^H]^H & \text{S1} \\ \hat{\mathbf{h}}_q & \text{S2} \\ [\hat{\mathbf{h}}_{q-1}^H \dots \hat{\mathbf{h}}_{q-K}^H]^H & \text{S3} \end{cases}, \quad (8)$$

where the $L^2 L_H \times 1$ vector $\hat{\mathbf{h}}_q$ denotes the estimate of \mathbf{h}_q . Moreover, $J(\mathbf{h}_q)$ in (7) is given by

$$J(\mathbf{h}_q) = \frac{1}{L} \text{Tr} \left[\alpha^2 \mathbf{P}^H \mathbf{G}_q^H \mathbf{G}_q \mathbf{P} + \alpha^2 \frac{\mathbf{R}_n}{\sigma_a^2} + \mathbf{T}^H \mathbf{T} - \alpha \mathbf{T}^H \mathbf{G}_{T,q} \mathbf{P} - \alpha \mathbf{P}^H \mathbf{G}_{T,q}^H \mathbf{T} \right] \quad (9)$$

with

$$\mathbf{G}_q = [\mathbf{G}_q^H(-L_G^{(1)}) \dots \mathbf{G}_q^H(L_G^{(2)})]^H, \quad (10)$$

$$\mathbf{G}_{T,q} = [\mathbf{G}_q^H(0) \dots \mathbf{G}_q^H(L_T)]^H, \quad (11)$$

$$\mathbf{T} = [\mathbf{T}^H(0) \cdots \mathbf{T}^H(L_T)]^H, \quad (12)$$

and $\mathbf{R}_n = \mathbb{E}[\mathbf{n}(k)\mathbf{n}^H(k)]$.

With the consideration that the TIR and the pre-equalizer are functions of $\hat{\mathbf{h}}_q$, the MSE (7) is minimized by selecting

$$(\mathbf{T}_q^*, \mathbf{P}_q^*, \alpha_q^*) = \arg \min_{\mathbf{T}, \mathbf{P}, \alpha} Q(\hat{\mathbf{h}}_q), \quad (13)$$

where $Q(\hat{\mathbf{h}}_q) = \mathbb{E}[J(\mathbf{h}_q)|\hat{\mathbf{h}}_q]$ is the a posteriori expectation of $J(\mathbf{h}_q)$, given the imperfect CSIT $\hat{\mathbf{h}}_q$. In the case of perfect CSIT, we have $\hat{\mathbf{h}}_q = \mathbf{h}_q$, so that (13) reduces to the minimization of $J(\mathbf{h}_q)$.

From (9) it follows that

$$Q(\hat{\mathbf{h}}_q) = \frac{1}{L} \text{Tr} \left[\alpha^2 \mathbf{P}^H \mathbf{E}_{\text{GG}}(\hat{\mathbf{h}}_q) \mathbf{P} + \alpha^2 \frac{\mathbf{R}_n}{\sigma_a^2} + \mathbf{T}^H \mathbf{T} - \alpha \mathbf{T}^H \mathbf{E}_{\text{GT}}(\hat{\mathbf{h}}_q) \mathbf{P} - \alpha \mathbf{P}^H \mathbf{E}_{\text{GT}}^H(\hat{\mathbf{h}}_q) \mathbf{T} \right], \quad (14)$$

where

$$\mathbf{E}_{\text{GG}}(\hat{\mathbf{h}}_q) = \mathbb{E}[\mathbf{G}_q^H \mathbf{G}_q | \hat{\mathbf{h}}_q] \quad (15)$$

and

$$\mathbf{E}_{\text{GT}}(\hat{\mathbf{h}}_q) = \mathbb{E}[\mathbf{G}_{\text{T},q} | \hat{\mathbf{h}}_q]. \quad (16)$$

Appendix A discusses how to analytically obtain the a posteriori expectations $\mathbf{E}_{\text{GG}}(\hat{\mathbf{h}}_q)$ and $\mathbf{E}_{\text{GT}}(\hat{\mathbf{h}}_q)$ for the considered CSIT scenarios in the case of MMSE channel estimation from pilot symbols.

B. Calculation of equalization and TIR parameters

As mentioned above, the optimal set of parameters $(\mathbf{P}_q^*, \alpha_q^*, \mathbf{T}_q^*)$ is obtained by minimizing $Q(\hat{\mathbf{h}}_q)$ from (14) for a given estimation vector $\hat{\mathbf{h}}_q$. First, the optimal pre-equalizer and scalar factor are derived for given \mathbf{T} . Substituting this solution into $Q(\hat{\mathbf{h}}_q)$ yields $Q_{\mathbf{P},\alpha}(\hat{\mathbf{h}}_q)$, which is further optimized over \mathbf{T} .

To obtain $(\mathbf{P}_q^*, \alpha_q^*)$ under the energy constraint (2) for given \mathbf{T} , the Lagrangian $\Lambda(\hat{\mathbf{h}}_q)$ with multiplier λ is introduced, i.e.,

$$\Lambda(\hat{\mathbf{h}}_q) = Q(\hat{\mathbf{h}}_q) + \lambda(\text{Tr}(\mathbf{P}^H \mathbf{P} - LE_{\text{TX}})). \quad (17)$$

At optimum, the multiplier λ is determined by equating to zero the derivatives of Λ with respect to \mathbf{P} and α , and is given by

$$\lambda^* = \alpha^2 \zeta, \quad (18)$$

where $\zeta = \frac{\sigma_n^2}{LE_{\text{TX}}}$. The corresponding \mathbf{P}_q^* and α_q^* are subsequently derived by evaluating the derivatives of the Lagrangian Λ with $\lambda = \lambda^*$. This approach yields

$$\mathbf{P}_q^* = \frac{1}{\alpha_q^*} \mathbf{D}^{-1} \mathbf{E}_{\text{GT}}^H(\hat{\mathbf{h}}_q) \mathbf{T} \quad (19)$$

and

$$\alpha_q^* = \sqrt{\frac{1}{LE_{\text{TX}}} \text{Tr}(\mathbf{T}^H \mathbf{E}_{\text{GT}}(\hat{\mathbf{h}}_q) \mathbf{D}^{-1} \mathbf{D}^{-1} \mathbf{E}_{\text{GT}}^H(\hat{\mathbf{h}}_q) \mathbf{T})} \quad (20)$$

with $\mathbf{D} = \mathbf{E}_{\text{GG}}(\hat{\mathbf{h}}_q) + \zeta \mathbf{I}$. After plugging (19) and (20) into (14), the optimized $Q_{\mathbf{P},\alpha}(\hat{\mathbf{h}}_q)$ is written as

$$Q_{\mathbf{P},\alpha}(\hat{\mathbf{h}}_q) = \text{Tr}(\mathbf{T}^H \mathbf{Z} \mathbf{T}) \quad (21)$$

with $\mathbf{Z} = \mathbf{I} - \mathbf{E}_{\text{GT}}(\hat{\mathbf{h}}_q) \mathbf{D}^{-1} \mathbf{E}_{\text{GT}}^H(\hat{\mathbf{h}}_q)$.

In the case of FRS, $L_T = 0$ and \mathbf{T} equals the identity matrix. In the case of PRS, the precoder \mathbf{T}_q^* that minimizes $Q_{\mathbf{P},\alpha}(\hat{\mathbf{h}}_q)$ from (21) must be determined with the additional constraint that all entries of \mathbf{T}_q^* are Gaussian integers and $\mathbf{T}_q^*(0)$ is a unimodular matrix. As the search over all feasible \mathbf{T} involves an infinite computational complexity, a suboptimal iterative algorithm is proposed instead, in which exactly one column of the TIR matrix is incremented in each iteration. More specifically, the TIR matrix in the i th iteration is denoted by \mathbf{T}_i and the l_i th column, $\mathbf{t}_{l_i,i}$, is incremented according to

$$\mathbf{t}_{l_i,i+1} = \mathbf{t}_{l_i,i} + \mathbf{t}_{\text{inc}}, \quad (22)$$

while all other columns remain unaltered. Since all elements of \mathbf{T} are restricted to the set of Gaussian integers, all elements of \mathbf{t}_{inc} must be Gaussian integer as well. Moreover, $|\det(\mathbf{T}_{i+1}(0))|$ remains equal to 1 when the first L elements of \mathbf{t}_{inc} are a linear combination of the first L elements of the other columns $\mathbf{t}_{l,i}$ with $l \neq l_i$. Therefore, \mathbf{t}_{inc} is decomposed as

$$\mathbf{t}_{\text{inc}} = [(\mathbf{t}_{\text{inc}}^{(0)})^H (\boldsymbol{\lambda}_{\text{inc}}^{(1)})^H]^H, \quad (23)$$

where $\mathbf{t}_{\text{inc}}^{(0)}$ contains the first L elements of \mathbf{t}_{inc} and $\boldsymbol{\lambda}_{\text{inc}}^{(1)}$ contains the remaining elements. Since $\mathbf{t}_{\text{inc}}^{(0)}$ must be a linear combination of columns from $\mathbf{T}_i(0)$, it is further decomposed as

$$\mathbf{t}_{\text{inc}}^{(0)} = \sum_{\substack{l=1 \\ l \neq l_i}}^L \lambda_{l,\text{inc}}^{(0)} \mathbf{t}_{l,i}^{(0)} = \mathbf{T}_{l_i,i}^{(0)} \boldsymbol{\lambda}_{\text{inc}}^{(0)}, \quad (24)$$

where $\boldsymbol{\lambda}_{\text{inc}}^{(0)}$ is the column vector containing the $L-1$ coefficients $\lambda_{l,\text{inc}}^{(0)}$, $\mathbf{t}_{l,i}^{(0)}$ is the l th column of $\mathbf{T}_i(0)$, and $\mathbf{T}_{l_i,i}^{(0)}$ is constructed by removing the l_i th column from $\mathbf{T}_i(0)$. Subsequently, when the Hermitian and positive-definite matrix \mathbf{Z} is factorized according to its Cholesky decomposition, i.e., $\mathbf{Z} = \mathbf{L}\mathbf{L}^H$ with \mathbf{L} lower triangular, $Q_{\mathbf{P},\alpha,i+1}(\hat{\mathbf{h}}_q)$ in iteration $i+1$ can be rewritten as

$$Q_{\mathbf{P},\alpha,i+1}(\hat{\mathbf{h}}_q) = \text{Tr}(\mathbf{T}_{i+1}^H \mathbf{L}\mathbf{L}^H \mathbf{T}_{i+1}) = \sum_{l=1}^L \|\mathbf{L}^H \mathbf{t}_{l,i+1}\|^2. \quad (25)$$

In (25), the i th iteration alters only the term corresponding to $l = l_i$ and, after plugging (22), (23), and (24) into (25), the optimal $\boldsymbol{\lambda}_{\text{inc}}^* = [(\boldsymbol{\lambda}_{\text{inc}}^{(0)})^H (\boldsymbol{\lambda}_{\text{inc}}^{(1)})^H]^H$ can then be found by solving the following lattice decoding problem

$$\boldsymbol{\lambda}_{\text{inc}}^* = \arg \min_{\boldsymbol{\lambda}_{\text{inc}}} \|\mathbf{G}_{\text{lat}} \boldsymbol{\lambda}_{\text{inc}} + \mathbf{x}\|^2, \quad (26)$$

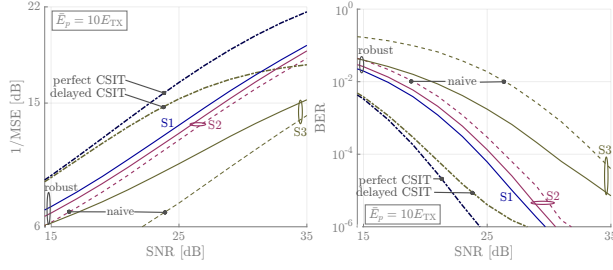


Figure 2. MSE (right plot) and BER (left plot) performance of optimized spatio-temporal PRS with $\bar{E}_p = 10E_{TX}$. The robust scenarios S1, S2, and S3 are compared to the cases of perfect CSIT and delayed CSIT ($K = 2$), and to the naive approach which ignores channel estimation errors. Clearly, the robust scenarios outperform their naive counterparts.

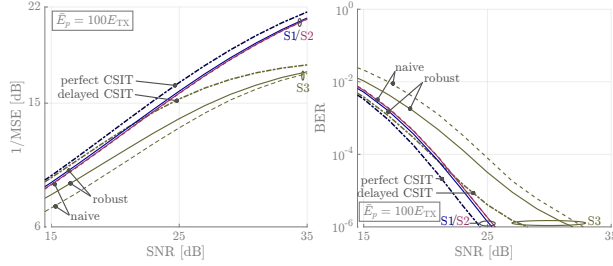


Figure 3. MSE (right plot) and BER (left plot) performance of optimized spatio-temporal PRS with $\bar{E}_p = 100E_{TX}$. The same scenarios as in Fig. 2 are compared. The increase of \bar{E}_p ensures that the performances for both the robust and naive approach get closer to the performances with perfect CSIT (scenarios S1 and S2) or delayed CSIT (scenario S3).

where $\mathbf{G}_{\text{lat}} = [\mathbf{L}^{(0)} \mathbf{T}_{l_i, i}^{(0)} \quad \mathbf{L}^{(1)}]$, $\mathbf{x} = -\mathbf{L}^H \mathbf{T}_{l_i, i}^{(0)}$, and $\mathbf{L}^{(0)}$ and $\mathbf{L}^{(1)}$ contain the first L and last LL_T columns of \mathbf{L}^H , respectively. To solve the minimization problem in (26), the closest point search algorithm described in [13] can readily be applied after the decomposition of all quantities into their real and imaginary parts. In each iteration, l_i corresponds to the largest term in the summation of (25) for which an increment can be found that reduces $Q_{\mathbf{P}, \alpha, i+1}(\hat{\mathbf{h}}_q)$. When none of the terms can be reduced, the algorithm terminates.

IV. NUMERICAL RESULTS AND DISCUSSION

To investigate the performance of the proposed robust MIMO PRS, this section presents numerical results pertaining to the transmission of precoded 4-QAM symbols over a FS Rayleigh fading channel at a symbol rate $1/T = 20$ MHz. The TX consists of (i) a PRS precoder with $L_T = 2$ feedback taps; (ii) a 21-tap ($L_P^{(1)} = L_P^{(2)} = 10$) pre-equalizer; and (iii) $L = 3$ antennas. The discrete-time channel impulse response matrix $\mathbf{H}_q(m)$ is related to the impulse response matrix $\mathbf{H}_{\text{ch}}^{(q)}(\tau)$ of the underlying continuous-time Rayleigh fading channel by $\mathbf{H}_q(m) = \int h_c(mT - \tau) \mathbf{H}_{\text{ch}}^{(q)}(\tau) d\tau$, where $h_c(u)$ is the impulse response of the cascade of the TX and RX filters; we take for $h_c(u)$ a cosine-rolloff pulse with

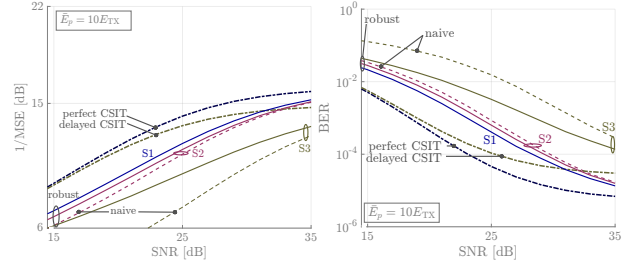


Figure 4. MSE (right plot) and BER (left plot) performance of FRS with $\bar{E}_p = 10E_{TX}$. The same scenarios as in Fig. 2 are compared. A significant performance degradation compared to optimized spatio-temporal PRS is observed.

30% rolloff. The elements of $\mathbf{H}_{\text{ch}}^{(q)}(\tau)$ are assumed to be independent zero-mean circular-symmetric Gaussian random variables, characterized by

$$\mathbb{E} \left[\left(\mathbf{H}_{\text{ch}}^{(q+k)}(\tau + u) \right)_{l_1, l_2} \left(\mathbf{H}_{\text{ch}}^{(q_2)}(\tau_2) \right)_{l_1, l_2}^* \right] = p_d(\tau) R_D(kXT) \delta(u), \quad (27)$$

where $p_d(\tau)$ is the power-delay profile (with $\int p_d(\tau) d\tau = 1$), and $R_D(w)$ is the inverse Fourier transform of the Doppler spectrum $p_D(\nu)$ (with ν the Doppler frequency and $\int p_D(\nu) d\nu = 1$). We take $p_d(\tau) = (1/T) \exp(-\tau/T)$. As for the Doppler spectrum $p_D(\nu)$, a zero-mean Gaussian with a standard deviation equal to $f_D = 50$ Hz is selected. Taking symbol blocks of $N_{\text{tot}} = 10^4$ symbols, we obtain $N_{\text{tot}} T f_D = 0.025 \ll 1$, indicating that the channel variations over a symbol block are small. The relation between $\mathbf{H}_q(m)$ and $\mathbf{H}_{\text{ch}}^{(q)}(\tau)$ and the property (27) allow to numerically compute the covariance between $(\mathbf{H}_{q_1}(m_1))_{l_1, l_2}$ and $(\mathbf{H}_{q_2}(m_2))_{l_1, l_2}$ for the relevant values of $m_1, m_2, q_1 - q_2$; these covariances do not depend on (l_1, l_2) . Channel estimation is accomplished using Zadoff-Chu (ZC) sequences, with $\bar{E}_p = 10E_{TX}$ or $\bar{E}_p = 100E_{TX}$ (see Appendix B).

Both the MSE performance and the BER performance are averaged over 2000 channel realizations and depicted in Figs. 2, 3 and 4 for the optimized PRS with $\bar{E}_p = 10$, the optimized PRS with $\bar{E}_p = 100$ and FRS with $\bar{E}_p = 10$, respectively. The signal-to-noise ratio (SNR) is defined as the ratio E_{TX}/σ_n^2 . In all figures, the proposed robust scenarios S1, S2, and S3 (with $K = 2$) are compared to the following cases:

- Perfect CSIT, where the actual channel \mathbf{h}_q is available. This case outperforms all other scenarios.
- Delayed CSIT, where the K previous channels are available and collected in the vector $\bar{\mathbf{h}}_{q, K} = [\mathbf{h}_{q-1}^H, \dots, \mathbf{h}_{q-K}^H]^H$. To determine the optimal equalizer and TIR, the a posteriori expectation $\mathbb{E}[\cdot | \hat{\mathbf{h}}_q]$ in (15) and (16) is replaced by the a posteriori expectation $\mathbb{E}[\cdot | \bar{\mathbf{h}}_{q, K}]$ such that $Q(\cdot)$ defined in (14) becomes a function of $\bar{\mathbf{h}}_{q, K}$ instead of $\hat{\mathbf{h}}_q$. The

associated conditional PDF can directly be derived from the channel statistics. The case of delayed CSIT outperforms scenario S3.

- The naive approach, where channel estimation errors are ignored. More precisely, in S2 the equalizer and TIR coefficients minimize $J(\mathbf{h}_q)$ from (9), but with \mathbf{h}_q replaced by the current estimate $\hat{\mathbf{h}}_q$. Similarly, in S3, $Q(\bar{\mathbf{h}}_{q,K})$ defined in the case of delayed CSIT is minimized, but with $\bar{\mathbf{h}}_{q,K}$ replaced by the delayed estimates $\hat{\mathbf{h}}_{q,K}$. Both cases are inferior to their robust counterparts.

An especially noteworthy observation is the significant performance gain of PRS compared to FRS. For the cases with perfect CSIT and delayed CSIT, PRS indeed achieves not only a considerably lower MSE than FRS in the high SNR region, but also significantly lowers the BER floor (on the order of 10^{-5} for FRS) to (far) below 10^{-6} , emphasizing the potential of spatio-temporal PRS. Likewise, all robust strategies in the case of PRS, including S3, not only outperform their robust counterparts in FRS, but also reach (considerably) lower BER values (at high SNR) than FRS with perfect CSIT, even for the rather poor estimation quality of $\bar{E}_p = 10E_{\text{TX}}$. This observation thus confirms that a suitable set of TIR and pre-equalizer is obtained by all robust PRS scenarios.

Not surprisingly, the more channel estimates available at the TX, the better the performance of the robust scenario becomes, which is illustrated by the superior performance of S1 and S2 compared to S2 and S3, respectively. When the estimation quality improves, e.g., $\bar{E}_p = 100E_{\text{TX}}$ in Fig. 3, the performance gap between robust S1 and robust S2 diminishes and both strategies get close to the case with perfect CSI, because an (excellent) estimate of the current channel is available. Although the MSE and BER of S3 reduces with increasing \bar{E}_p , a performance loss compared to S1/S2 is noticeable for both PRS and FRS; a similar observation holds when comparing perfect CSIT with delayed CSIT (which are the performance limits for S1/S2 and S3 when \bar{E}_p grows infinitely large). This performance gap between S1/S2 and S3 is due to the rather large uncertainty about the current channel given the estimates of only $K = 2$ prior channel realizations, and verified to nearly disappear for increasing K (results not shown for conciseness).

Figs. 2 and 4 reveal that the difference between the performances of the robust strategies (S1/S2 and S3) and their limits for infinite \bar{E}_p (perfect CSIT and delayed CSIT) is larger for PRS than for FRS. This is caused by the additional optimization over the TIR in PRS. Indeed, whereas the TIR for FRS is the same for all strategies, for PRS the robust TIR is not necessarily identical to the TIR derived with perfect CSIT (S1/S2) or delayed CSIT (S3); these TIR differences could induce a considerable difference in both MSE and BER, since both the real and imaginary part of the TIR are not continuous because

they take integer values only.

In the naive approach, the equalization and TIR parameters are computed assuming that the current estimate $\hat{\mathbf{h}}_q$ (S2) or the previous estimates $\hat{\mathbf{h}}_{q,K}$ (S3) are without estimation error. Figs. (2) - (4) clearly demonstrate that the naive approach is inferior to the robust approach. The difference between the robust and naive approach is particularly prominent when the channel estimates are rather poor, whereas the difference between the two approaches reduces, as expected, with increasing \bar{E}_p . Hence, the robust approach is preferable, especially since the complexity of both approaches is comparable as it is dominated by the computation of \mathbf{T} and not by the evaluation of the expectations in (15) and (16).

V. CONCLUSION

This paper applies the concept of robust equalization to spatio-temporal PRS over a FS and time-variant MIMO channel. More precisely, the TX designs both the TIR of the PRS and the linear pre-equalizer, based solely on the available imperfect CSIT, which consists of noisy estimates of the MIMO channel of the current and/or previous blocks. The robustness in the design is accomplished by statistically modeling the error of the estimates and subsequently minimizing the MSE between the decision variable and the target response; this MSE incorporates the joint PDF of the actual channel realization and the available CSIT. Numerical results confirm that the proposed robust PRS significantly improves the performance compared to the naive approach that ignores channel estimation errors. Compared to FRS, PRS also improves the MSE and BER performance for both the robust design and the design with perfect CSIT, making the proposed robust PRS an interesting technique to handle considerable amounts of ISI in practice.

APPENDIX A

LINEAR MMSE ESTIMATION FROM PILOT SYMBOLS

During the q th symbol block, each antenna transmits N_p pilot symbols (besides the data symbols) to allow channel estimation at the RX. To avoid that the channel dispersion gives rise to interference from the pilot symbols to the data symbols in the same block, the TX inserts a guard interval of $L_H - 1$ all-zero symbols between the pilot symbols and the data symbols¹. Thus, the block size N_{tot} is given by $N_p + N_d + L_H - 1$, where N_d denotes the number of data symbols per block.

Only the signals received at the $N_p - L_H + 1$ consecutive instants, where the channel memory is entirely filled with pilot symbols, are exploited for channel estimation. This yields the $L \times (N_p - L_H + 1)$ observation matrix $\mathbf{Y}_q = \mathbf{H}_q \mathbf{A}_p + \mathbf{N}_q$, where the i th column of the $LL_H \times (N_p - L_H + 1)$ pilot symbol matrix \mathbf{A}_p

¹As the channel covariance matrix is assumed to be known, the value of L_H can be derived

equals $(\mathbf{a}_p^H(L_H + i - 2), \dots, \mathbf{a}_p^H(i - 1))^H$, with $\mathbf{a}_p(n)$ the $L \times 1$ pilot symbol vector transmitted during the n th symbol interval in the q th block. This observation can equivalently be represented as $\mathbf{y}_q = \mathbf{A}\mathbf{h}_q + \mathbf{n}_q$, where $\mathbf{A} = \mathbf{A}_p^T \otimes \mathbf{I}_L$ and \mathbf{y}_q , \mathbf{n}_q , and \mathbf{h}_q are obtained by stacking the columns of \mathbf{Y}_q , \mathbf{N}_q , and \mathbf{H}_q , respectively. The covariance matrix of the noise contribution \mathbf{n}_q is given by $\mathbf{R}_n = \sigma_n^2 \mathbf{I}_{L(N_p - L_H + 1)}$. The resulting linear MMSE estimate of \mathbf{h}_q is $\hat{\mathbf{h}}_q = \mathbf{X}\mathbf{y}_q$ with

$$\mathbf{X} = \left(\sigma_n^2 \mathbf{R}_{\mathbf{h}_q}^{-1} + \mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{A}^H, \quad (28)$$

where $\mathbf{R}_{\mathbf{h}_q}$ is the covariance matrix of \mathbf{h}_q . Introducing the $L_H \times L_H$ matrices $\mathbf{U}(i)$, where

$$(\mathbf{U}(i))_{m,m'} = \mathbb{E} \left[(\mathbf{H}_{q+i}(m))_{l,l'} (\mathbf{H}_q(m'))_{l,l'}^* \right] \quad (29)$$

for $(m, m') \in \{-L_H^{(1)}, \dots, L_H^{(2)}\}^2$, we obtain $\mathbf{R}_{\mathbf{h}_q} = \mathbf{U}(0) \otimes \mathbf{I}_{L^2}$. Defining $\mathbf{h}_q = \mathbf{h}_q$ for CSIT scenario S2, and $\bar{\mathbf{h}}_q = [\mathbf{h}_q^H \dots \mathbf{h}_{q-K}^H]^H$ for scenarios S1 and S3, it follows from the decomposition $\hat{\mathbf{h}}_q = \mathbf{X}\mathbf{A}\mathbf{h}_q + \mathbf{X}\mathbf{n}_q$ that $\hat{\mathbf{h}}_q = \mathbf{E}\bar{\mathbf{h}}_q + \mathbf{F}\bar{\mathbf{n}}_q$, where

$$\mathbf{E} = \begin{cases} \mathbf{I}_{K+1} \otimes (\mathbf{X}\mathbf{A}) & \text{S1} \\ \mathbf{X}\mathbf{A} & \text{S2} \\ [\mathbf{0} \ \mathbf{I}_K \otimes (\mathbf{X}\mathbf{A})] & \text{S3} \end{cases} \quad (30)$$

and \mathbf{F} is obtained by replacing in (30) $\mathbf{X}\mathbf{A}$ by \mathbf{X} , and $\bar{\mathbf{n}}_q$ results from stacking the vectors $\mathbf{n}_q, \dots, \mathbf{n}_{q-K}$. Hence, conditioned on $\bar{\mathbf{h}}_q$, $\hat{\mathbf{h}}_q$ is Gaussian with mean $\mathbf{E}\bar{\mathbf{h}}_q$ and covariance matrix $\sigma_n^2 \mathbf{F}\mathbf{F}^H$. The vector $\bar{\mathbf{h}}_q$ is zero-mean Gaussian with covariance matrix $\mathbf{R}_{\bar{\mathbf{h}}_q} = \bar{\mathbf{U}} \otimes \mathbf{I}_{L^2}$, where $\bar{\mathbf{U}}$ is a $(K+1)L_H \times (K+1)L_H$ block matrix, with the $L_H \times L_H$ block at position $(i, j) \in \{1, \dots, K+1\}^2$ given by $\mathbf{U}(i-j)$. It can be verified that, conditioned on $\hat{\mathbf{h}}_q$, $\bar{\mathbf{h}}_q$ is Gaussian with mean $\mathbf{R}_{\bar{\mathbf{h}}_q|\hat{\mathbf{h}}_q} \mathbf{E}^H (\sigma_n^2 \mathbf{F}\mathbf{F}^H)^{-1} \hat{\mathbf{h}}_q$ and covariance matrix $\mathbf{R}_{\bar{\mathbf{h}}_q|\hat{\mathbf{h}}_q}$, where

$$\mathbf{R}_{\bar{\mathbf{h}}_q|\hat{\mathbf{h}}_q} = \left(\mathbf{R}_{\bar{\mathbf{h}}_q}^{-1} + \mathbf{E}^H (\sigma_n^2 \mathbf{F}\mathbf{F}^H)^{-1} \mathbf{E} \right)^{-1} \quad (31)$$

Based on (10) and (11), it follows that the elements of $\mathbf{E}_{\text{GG}}(\hat{\mathbf{h}}_q)$ and \mathbf{E}_{GT} from (15) and (16) can be straightforwardly expressed in terms of the a posteriori expectations and covariances of the elements of \mathbf{h}_q ; these a posteriori moments are obtained by extracting the proper elements from the a posteriori mean and the a posteriori covariance matrix of $\hat{\mathbf{h}}_q$.

APPENDIX B

ORTHOGONAL PILOT SEQUENCES

As an example, the considered pilot sequence is derived from a ZC sequence [14], [15]. The pilot sequence for the first antenna is constructed by cyclically extending a ZC sequence of length $N_p - L_H + 1$ with $L_H - 1$ symbols, which yields a pilot sequence of length N_p . The pilot sequence for the l th antenna is obtained

by applying $(l-1)L_H$ cyclic shifts to the ZC sequence from the first antenna, and adding a cyclic extension of $L_H - 1$ symbols. For $N_p - L_H + 1 \geq L_H L$, all rows from \mathbf{A}_p are orthogonal, yielding $\mathbf{A}_p \mathbf{A}_p^H = \bar{E}_p \mathbf{I}_{L L_H}$, where $\bar{E}_p = (N_p - L_H + 1)E_p$ and E_p is the transmitted energy per pilot symbol.

ACKNOWLEDGMENT

Part of this research has been supported by the Research Foundation Flanders (FWO) under Grant no. EOS-30458693 'MUSE-WINET) MUlti-Service Wireless NETwork' and Grant no. G.013917N 'Advanced MIMO channel modeling and signal processing for high-bitrate chip-to-chip interconnects prone to manufacturing variability'. Jelle Bailleul is supported by an UGent-BOF PhD scholarship.

REFERENCES

- [1] G. Forney, "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.
- [2] N. Al-Dhahir and A. H. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921–2936, Oct. 2000.
- [3] M. Joham, D. A. Schmidt, J. Brehmer, and W. Utschick, "Finite-Length MMSE Tomlinson-Harashima Precoding for Frequency Selective Vector Channels," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 3073–3088, Jun. 2007.
- [4] P. Kabal and S. Pasupathy, "Partial-Response Signaling," *IEEE Trans. Commun.*, vol. 23, no. 9, pp. 921–934, Sep. 1975.
- [5] L. Jacobs, J. Bailleul, P. Manfredi, M. Guenach, D. Vande Ginste, and M. Moeneclaey, "On Partial Response Signaling for MIMO Equalization on Multi-Gbit/s Electrical Interconnects," in *26th European Signal Processing Conference (EUSIPCO 2018)*, Sep. 2018.
- [6] R. F. H. Fischer and C. Siegl, "On the Relation between Lattice-Reduction-Aided Equalization and Partial-Response Signaling," in *2006 Int. Zurich Seminar on Communications*, 2006, pp. 34–37.
- [7] J. Bailleul, L. Jacobs, and M. Moeneclaey, "Optimized pre-coded spatio-temporal partial-response signaling over frequency-selective MIMO channels," Submitted to *IEEE Trans. on Wireless Commun.*
- [8] N. Vucic, H. Boche, and S. Shi, "Robust Transceiver Optimization in Downlink Multiuser MIMO Systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.
- [9] J. Wang and D. P. Palomar, "Worst-Case Robust MIMO Transmission With Imperfect Channel Knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086–3100, Aug. 2009.
- [10] T.-J. Ho and B.-S. Chen, "Robust Minimax MSE Equalizer Designs for MIMO Wireless Communications With Time-Varying Channel Uncertainties," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5835–5844, Nov. 2010.
- [11] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically Robust Design of Linear MIMO Transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.
- [12] F. A. Dietrich, P. Breun, and W. Utschick, "Robust Tomlinson-Harashima Precoding for the Wireless Broadcast Channel," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 631–644, Feb. 2007.
- [13] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest Point Search in Lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.
- [14] D. Chu, "Polyphase codes with good periodic correlation properties (corresp.)," *IEEE Trans. Inform. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
- [15] R. Frank, S. Zadoff, and R. Heimiller, "Phase shift pulse codes with good periodic correlation properties (corresp.)," *IRE Transactions on Information Theory*, vol. 8, no. 6, pp. 381–382, Oct. 1962.